arXiv:1209.3691v1 [math.PR] 17 Sep 2012

An old approach to the giant component problem

Béla Bollobás*† Oliver Riordan‡

September 17, 2012

Abstract

In 1998, Molloy and Reed showed that, under suitable conditions, if a sequence $\mathbf{d}_n$ of degree sequences converges to a probability distribution D, then the size of the largest component in the n-vertex random graph associated to $\mathbf{d}_n$ is asymptotically $\rho(D)n$, where $\rho(D)$ is a constant defined by the solution to certain equations that can be interpreted as the survival probability of a branching process associated to D. There have been a number of papers strengthening this result in various ways; here we prove a strong form of the result (with exponential bounds on the probability of large deviations) under minimal conditions.

1 Introduction and results

By a *degree sequence* $\mathbf{d}$ we mean a finite sequence $(d_1, \ldots, d_n)$ of non-negative integers with even sum. The *length* $|\mathbf{d}|$ of $\mathbf{d} = (d_i)_{i=1}^n$ is the number n of terms, and the *size* $m(\mathbf{d}) = \frac{1}{2}\sum_i d_i$ is half the sum of the terms. We write $G_{\mathbf{d}}$ for the *random (simple) graph with degree sequence* $\mathbf{d}$, i.e., a graph with vertex set $[n] = \{1, 2, \ldots, n\}$ in which each vertex i has degree d_i, chosen uniformly at random from the set of all such graphs (assuming this set is non-empty). As usual, in studying $G_{\mathbf{d}}$ we also consider the corresponding random *configuration multigraph* $G^*_{\mathbf{d}}$, introduced in [2], obtained by associating a set of d_i stubs to each vertex i, selecting a uniformly random pairing of the (disjoint) union of these sets, and interpreting each paired pair of stubs as leading to an edge in the natural way. Note that these graphs have $|\mathbf{d}|$ vertices and $m(\mathbf{d})$ edges.

Let $\mathcal{D}$ denote the set of probability distributions D on the non-negative integers with $0 < \mathbb{E}(D) < \infty$. We usually write $D \in \mathcal{D}$ as $D = (r_0, r_1, \ldots)$, where, abusing notation by also writing D for a random variable with distribution D, $r_i = \mathbb{P}(D = i)$. One of the basic questions concerning the random graph models

*Department of Pure Mathematics and Mathematical Statistics, Wilberforce Road, Cambridge CB3 0WB, UK and Department of Mathematical Sciences, University of Memphis, Memphis TN 38152, USA. E-mail: `b.bollobas@dpmms.cam.ac.uk`.

†Research supported in part by NSF grant DMS-0906634

‡Mathematical Institute, University of Oxford, 24–29 St Giles', Oxford OX1 3LB, UK. E-mail: `riordan@maths.ox.ac.uk`.

just described is the following: under what conditions does convergence of $\mathbf{d}_n$ to D imply that the asymptotic behaviour of $G_{\mathbf{d}_n}$ (or $G^*_{\mathbf{d}_n}$) is captured by D? Here the behaviour we are interested in is the distribution of the component sizes, and most particularly the number L_1 of vertices in the (a if there is a tie) largest component.

Let

$$n_i(\mathbf{d}) = \big|\{j : d_j = i\}\big|$$

denote the number of times a particular degree i occurs in $\mathbf{d}$, so

$$m(\mathbf{d}) = \frac{1}{2}\sum_{j=1}^{|\mathbf{d}|} d_j = \frac{1}{2}\sum_{i=0}^{\infty} i n_i(\mathbf{d}).$$

The basic assumptions made in all existing results of this type are that

$$\lim_{n\to\infty} \frac{n_i(\mathbf{d}_n)}{|\mathbf{d}_n|} = r_i \tag{1}$$

for each i, that

$$\frac{m(\mathbf{d}_n)}{|\mathbf{d}_n|} \to \frac{\mathbb{E}(D)}{2} = \frac{1}{2}\sum_{i=0}^{\infty} i r_i \tag{2}$$

as $n \to \infty$, and of course that $|\mathbf{d}_n| \to \infty$. (Often, one takes $|\mathbf{d}_n| = n$, which loses no generality.) We shall say that $\mathbf{d}_n$ *converges* to D, and write $\mathbf{d}_n \to D$, if these conditions hold.

Condition (1) ensures that D captures the asymptotic proportion of vertices of each fixed degree, and condition (2) that the (rescaled) number of edges is related to D in the natural way. Note that if we write D_n for the distribution of a randomly chosen element of $\mathbf{d}_n$ (i.e., the degree of a random vertex of $G_{\mathbf{d}_n}$), then (1) asserts that D_n converges to D in distribution. Condition (2) asserts that $\mathbb{E}(D_n) \to \mathbb{E}(D) < \infty$, which (given (1)) is equivalent to uniform integrability of the D_n.

To see why (2) is necessary, consider $\mathbf{d}_n$ consisting of one vertex of degree $n-1$ and $n-1$ of degree 1, contrasted (for n even) with $\mathbf{d}'_n$ in which all n degrees are equal to 1. In both cases (1) holds with $r_1 = 1$ and all other $r_i = 0$, but the component structures of $G_{\mathbf{d}_n}$ and $G_{\mathbf{d}'_n}$ are very different – one is a star, and the other a matching. (There is a similar but less extreme difference between $G^*_{\mathbf{d}_n}$ and $G^*_{\mathbf{d}'_n}$.)

As usual, we write $L_i(G)$ for the number of vertices in the ith largest component of a graph G. We also write $N_k(G)$ for the number of vertices in k-vertex components. The next result involves constants $\rho_k(D)$ and $\rho(D)$ whose definitions we postpone to Section 2 (see (10)). In fact, $\rho(D)$ is the same as the quantity ε_D appearing in [12], although our definition of it is different. We write $\xrightarrow{\mathrm{p}}$ to denote convergence in probability.

Theorem 1. *Let D be a probability distribution on the non-negative integers with $0 < \mathbb{E}(D) < \infty$, and let $\mathbf{d}_n$ be a sequence of degree sequences converging to D in the sense that (1) and (2) hold and $|\mathbf{d}_n| \to \infty$. Then*

$$N_k(G_{\mathbf{d}_n})/|\mathbf{d}_n| \xrightarrow{\mathrm{p}} \rho_k(D)$$

for each fixed k. If $\mathbb{P}(D \geqslant 3) > 0$, then we also have

$$L_1(G_{\mathbf{d}_n})/|\mathbf{d}_n| \xrightarrow{\mathrm{p}} \rho(D)$$

and $L_2(G_{\mathbf{d}_n})/|\mathbf{d}_n| \xrightarrow{\mathrm{p}} 0$.

Furthermore, the same conclusions hold with $G_{\mathbf{d}_n}$ replaced by $G^*_{\mathbf{d}_n}$.

The first result of this type was proved by Molloy and Reed [12], but with additional conditions. Taking $|\mathbf{d}_n| = n$, they assumed in particular that the maximum degree in $\mathbf{d}_n$ is $o(n^{1/4-\varepsilon})$ for some $\varepsilon > 0$. Note that (2) implies only that the maximum degree is $o(n)$: adding a single vertex with degree (approximately) $n/\log\log n$, say, does not affect convergence in our sense.

The results of [12] have been strengthened in a number of ways. One main direction is to improve, or even study the distribution of, the error term in the result $L_1 = \rho(D)n + o_{\mathrm{p}}(n)$, sometimes imposing extra assumptions; see Kang and Seierstad [10], Pittel [13], Janson and Luczak [9], Riordan [15] and Hatami and Molloy [7], for example. In the other direction, one can ask for the same conclusion but with less restrictive assumptions; here Janson and Luczak [9] prove a version of Theorem 1 with the condition that the sum of the squares of the degrees is at most a constant times the number of vertices. They also prove the (easier) *multigraph* part of Theorem 1 under exactly our conditions (see their Remark 2.6), but using a very different method.

We shall in fact prove a much stronger form of Theorem 1, Theorem 2 below; the reason for postponing the statement is that it is a little more awkward: instead of convergence, we need to work with neighbourhoods. Given $D \in \mathcal{D}$ and a degree sequence $\mathbf{d}$, writing $r_i = \mathbb{P}(D = i)$ as usual, set

$$d^0_{\mathrm{conf}}(\mathbf{d}, D) = \sum_{i=1}^{\infty} \left| i\frac{n_i(\mathbf{d})}{|\mathbf{d}|} - ir_i \right|, \tag{3}$$

so d^0_{conf} is a form of the ℓ_1 metric, and define the *configuration distance* between $\mathbf{d}$ and D to be

$$d_{\mathrm{conf}}(\mathbf{d}, D) = \max\{d^0_{\mathrm{conf}}(\mathbf{d}, D), 1/|\mathbf{d}|\}. \tag{4}$$

The $1/|\mathbf{d}|$ term in (4) ensures that $d_{\mathrm{conf}}(\mathbf{d}, D) \to 0$ if and only if $d^0_{\mathrm{conf}}(\mathbf{d}, D) \to 0$ and $|\mathbf{d}| \to \infty$, and avoids having to write 'and $|\mathbf{d}| \geqslant n_0$' in many results below; this is a convenience rather than an essential part of the definition.

It is easy to check that, for $D \in \mathcal{D}$,

$$\mathbf{d}_n \to D \iff d_{\mathrm{conf}}(\mathbf{d}_n, D) \to 0. \tag{5}$$

For the reverse implication, if $d_{\mathrm{conf}}(\mathbf{d}_n, D) \to 0$, then certainly $|\mathbf{d}_n| \to \infty$. Also, $d^0_{\mathrm{conf}}(\mathbf{d}_n, D) \to 0$, which trivially implies (1), and implies (2) by the triangle inequality. For the forward direction, given any $\varepsilon > 0$, since $\mathbb{E}(D) < \infty$ there is some $C = C(\varepsilon)$ such that $\sum_{i \geqslant C} ir_i < \varepsilon/4$, say. For n large enough, (1) gives $|n_i(\mathbf{d}_n)/|\mathbf{d}_n| - r_i| < \varepsilon/(4C)$ for all $i < C$, and it follows that $\sum_{i<C} in_i(\mathbf{d}_n)/|\mathbf{d}_n|$ is within $\varepsilon/4$ of $\sum_{i<C} ir_i \geqslant \mathbb{E}(D) - \varepsilon/4$. Using (2) it follows that $\sum_{i \geqslant C} in_i(\mathbf{d}_n)/|\mathbf{d}_n| \leqslant 3\varepsilon/4$ if n is large, and hence that $d^0_{\mathrm{conf}}(\mathbf{d}_n, D) \leqslant \varepsilon$. Since $\mathbf{d}_n \to D$ implies $|\mathbf{d}_n| \to \infty$ by definition, it follows that $d_{\mathrm{conf}}(\mathbf{d}_n, D) \to 0$.

Let us state for future reference a consequence of the argument just given: if $\mathbf{d}_n \to D$ then

$$\forall \varepsilon > 0 \ \exists C \ \forall n \ \sum_{i \geqslant C} in_i(\mathbf{d}_n) \leqslant \varepsilon |\mathbf{d}_n|. \tag{6}$$

Writing $\mathbf{d}_n = (d^{(n)}_1, \ldots, d^{(n)}_{\ell_n})$, (6) can be written as

$$\forall \varepsilon > 0 \ \exists C \ \forall n \sum_{j \,:\, d^{(n)}_j \geqslant C} d^{(n)}_j \leqslant \varepsilon |\mathbf{d}_n|.$$

Informally, this condition says that a random edge has only a small probability of being attached to a vertex of very high degree. A rather trivial consequence of (6) is that, writing $\Delta(\mathbf{d})$ for the maximum degree appearing in a degree sequence $\mathbf{d}$, if $\mathbf{d}_n \to D$ then $\Delta(\mathbf{d}_n) = o(|\mathbf{d}_n|)$. In terms of the metric, the equivalent of (6) is the observation that

$$\forall D \in \mathcal{D}, \varepsilon > 0 \ \exists C, \delta : d_{\mathrm{conf}}(\mathbf{d}, D) < \delta \Longrightarrow \sum_{i \geqslant C} in_i(\mathbf{d}) \leqslant \varepsilon |\mathbf{d}|. \tag{7}$$

To see this, simply choose C such that $\sum_{i \geqslant C} i\mathbb{P}(D = i) < \varepsilon/2$, and take $\delta = \varepsilon/2$.

Theorem 2. *Let $D \in \mathcal{D}$, and let $\varepsilon > 0$. For each $k \geqslant 1$ there exists $\delta > 0$ (depending on D, ε and k) such that if $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$, then*

$$\mathbb{P}\big(|N_k(G_{\mathbf{d}}) - \rho_k(D)n| \geqslant \varepsilon n\big) \leqslant e^{-\delta n}, \tag{8}$$

where $n = |\mathbf{d}|$. Moreover, if $\mathbb{P}(D \geqslant 3) > 0$, then there exists $\delta > 0$ (depending on D and ε) such that if $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$ then

$$\mathbb{P}\Big(\big|L_1(G_{\mathbf{d}}) - \rho(D)n\big| \geqslant \varepsilon n\Big) \leqslant e^{-\delta n}$$

and

$$\mathbb{P}\big(L_2(G_{\mathbf{d}}) \geqslant \varepsilon n\big) \leqslant e^{-\delta n}.$$

Furthermore, the same conclusions hold if $G_{\mathbf{d}}$ is replaced by $G^*_{\mathbf{d}}$.

Using (5), it is easy to check that Theorem 2 does indeed strengthen Theorem 1. The main reason for proving the stronger bounds in Theorem 2 is that we need them for the configuration multigraph model $G^*_{\mathbf{d}}$ in order to prove even Theorem 1 for the simple random graph $G_{\mathbf{d}}$. Of course, they are also nice to have!

Remark 3. The condition $\mathbb{P}(D \geqslant 3) > 0$ in Theorems 1 and 2 is necessary for the conclusions; see Janson and Luczak [9, Remark 2.7] for a discussion of the range of possible behaviours when $\mathbb{P}(D = 2) = 1$ (or D is supported on $\{0, 2\}$).

The basic idea of the proof of Theorem 1 is to use a (relatively) old method. The first ingredient is to understand the local structure of $G^*_{\mathbf{d}}$; this is very simple and can be expressed in a number of ways, most cleanly by comparison with a branching process. This allows us to control the number of vertices in small components. Then we use a version of the original sprinkling argument of Erdős and Rényi [4] to show that almost all vertices in 'large' components are in a single giant component. Of course, sprinkling is more complicated in the present model than in the original context. Also, to obtain exponential error bounds we need a strong form of the branching process approximation, which introduces some additional complications. We shall show in Section 6 that this approximation carries over to the giant component: the number of vertices in the giant component with some 'local' property can be calculated in terms of the branching process.

Turning to the nitty-gritty, in the rest of the paper we use the following standard asymptotic notation: given a sequence E_n of events, we say that E_n holds *with high probability* or *whp* if $\mathbb{P}(E_n) \to 1$ as $n \to \infty$. Given functions f and g of some parameter (usually n), $f = O(g)$ means f is bounded by a constant times g, and $f = o(g)$ means that $f/g \to 0$ as the parameter (n) tends to infinity.

Finally, before turning to the proofs, let us fix our formal notation for the configuration model: given a degree sequence $\mathbf{d}$ of length ℓ, we take disjoint sets $F_1, \ldots, F_\ell$ with $|F_i| = d_i$, where F_i represents the 'stubs' associated to vertex i. Then we take a pairing (partition into 2-elements sets) π of $F = \bigcup_{i=1}^{\ell} F_i$ chosen uniformly at random, and set $G^*_{\mathbf{d}} = \phi_{\mathbf{d}}(\pi)$, where $\phi_{\mathbf{d}}$ maps a pairing π to a multi-graph on $[\ell] = \{1, 2, \ldots, \ell\}$ by replacing each pair $\{a, b\}$ by an edge joining vertices i and j where $a \in F_i$ and $b \in F_j$, noting that $i = j$ is possible, in which case the edge is a loop.

2 Local approximation by a branching process

Let $D = (r_0, r_1, \ldots) \in \mathcal{D}$, so D is a probability distribution on the non-negative integers with $0 < \mathbb{E}(D) < \infty$, and $r_i = \mathbb{P}(D = i)$. For $i \geqslant 1$ let

$$q_i = \frac{i r_i}{\sum_i i r_i} = \frac{i r_i}{\mathbb{E}(D)}.$$

The distribution D^* with $\mathbb{P}(D^* = i) = q_i$ is known as the *size-biased* distribution associated to D. In any graph G, if we pick a random edge and then choose one of its endvertices v at random, the distribution of the degree of v is the size-biased version of the degree distribution of G. Let $Z = D^* - 1$, so

$$\mathbb{P}(Z = i) = \mathbb{P}(D^* = i + 1) = \frac{(i+1) r_{i+1}}{\mathbb{E}(D)}. \tag{9}$$

Intuitively, Z will correspond to the number of 'new' edges we get to when we follow a random edge to an endvertex.

Let $\mathcal{T}^1 = \mathcal{T}^1_D$ be the Galton–Watson branching process with offspring distribution Z, so $\mathcal{T}^1$ is a random rooted tree in which the number of children of each vertex has distribution Z, with these numbers independent. Finally, let $\mathcal{T} = \mathcal{T}_D$ be the random rooted tree in which the degree of the root has the distribution D, and, given the degree of the root, the branches, i.e., the subtrees rooted at the children of the root, form independent copies of $\mathcal{T}^1$.

It is not hard to see that if $\mathbf{d}_n \to D$, then $G^*_{\mathbf{d}_n}$ 'locally looks like' $\mathcal{T}_D$; we shall make this precise in a moment. Let $|\mathcal{T}_D| \leqslant \infty$ denote the total number of vertices of $\mathcal{T}_D$. Then the constants ρ_k and ρ appearing in Theorems 1 and 2 are

$$\rho_k(D) = \mathbb{P}(|\mathcal{T}_D| = k) \quad \text{and} \quad \rho(D) = \mathbb{P}(|\mathcal{T}_D| = \infty). \tag{10}$$

Given a graph G, for $v \in V(G)$ and $t \geqslant 0$, let $\Gamma_{\leqslant t}(v) = \Gamma^G_{\leqslant t}(v)$ denote the subgraph of G induced by the vertices within distance t of v, regarded as a rooted graph with root v. Also, let $\mathcal{T}_D|_t$ be the subtree of $\mathcal{T}_D$ induced by the vertices within distance t of the root. The following lemma is a variant of an idea that is by now very much standard, though perhaps not in exactly this form.

Lemma 4. *Let $D \in \mathcal{D}$ and suppose that $\mathbf{d}_n \to D$. Let v be a vertex of $G = G^*_{\mathbf{d}_n}$ chosen uniformly at random. Then we may couple the random graphs $\Gamma^G_{\leqslant t}(v)$ and $\mathcal{T}_D|_t$ so that they are isomorphic as rooted graphs with probability $1 - o(1)$ as $n \to \infty$.*

Proof. As the argument is straightforward and standard we give only an outline. The idea is to reveal $\Gamma_{\leqslant t}(v)$ step-by-step in the natural way, coupling this process with revealing $\mathcal{T}_D|_t$ step-by-step so that for any fixed j, the probability of the coupling failing at step j is $o(1)$. Since, given any ε, there is some constant J such that with probability at least $1 - \varepsilon$ the finite tree $\mathcal{T}_D|_t$ has size at most J, this suffices to prove the lemma.

To reveal $\Gamma_{\leqslant t}(v)$, first pick the random vertex v, noting that by condition (1) of the convergence $\mathbf{d}_n \to D$, the degree of v can be coupled to agree with the degree of the root of $\mathcal{T}_D$ with probability $1 - o(1)$. Then go through the stubs associated to v one-by-one, revealing their partners, and thus the neighbours of v (as well as any loops at v). Then reveal the partners of the unpaired stubs associated to the neighbours of v, and so on. The key fact is that the jth time we reveal the partner of an unpaired stub, the probability that this is a 'new' (not so far reached in the exploration) vertex of degree i is exactly

$$\frac{i(n_i(\mathbf{d}_n) - u_{i,j})}{2m(\mathbf{d}_n) + 1 - 2j},$$

where $u_{i,j}$ is the number of degree-i vertices revealed so far. For any fixed j, since $u_{i,j} \leqslant j = O(1)$, this probability is $q_i + o(1)$. Since q_i is the probability that a vertex of $\mathcal{T}_D$ other than the root has degree i (and hence $i - 1$ children), it follows that the coupling succeeds at step j with probability $1 - o(1)$, as required. $\square$

Corollary 5. *Let $D \in \mathcal{D}$, suppose that $\mathbf{d}_n \to D$, and let $t \geqslant 1$ be constant. Let v be a vertex of $G = G^*_{\mathbf{d}_n}$ chosen uniformly at random. Then whp the neighbourhood $\Gamma_{\leqslant t}(v)$ of v in G is a tree.* □

Note that in many related situations, the equivalent of Corollary 5 is proved by considering the expected number of paths of length k ending in a vertex on a cycle of length ℓ, showing that this expectation is $o(n)$ for k and ℓ fixed. However, this requires some condition such as $\sum_i d_i^2 = n^{1+o(1)}$, which need not hold here – it may be that $G^*_{\mathbf{d}}$ contains many (more than n) short cycles, but these are all concentrated in the neighbourhoods of the few vertices with largest degrees, so most vertices are far from them.

Let $\mathcal{P}$ be a property of (locally finite) rooted graphs, i.e., a set of rooted graphs closed under isomorphism. Often we think of $\mathcal{P}$ as a property of vertices v of unrooted graphs G, by taking v as the root; in either case we write $(G, v) \in \mathcal{P}$ to mean that the graph G rooted at v has property $\mathcal{P}$. We write $N_{\mathcal{P}}(G)$ for the number of vertices of G with property $\mathcal{P}$. Given $t \geqslant 1$, we say that $\mathcal{P}$ is *t-local* if whether (G, v) has $\mathcal{P}$ depends only on the rooted graph $\Gamma^G_{\leqslant t}(v)$. We call $\mathcal{P}$ *local* if it is t-local for some t. Note that it makes sense to speak of our branching process $\mathcal{T}_D$ having property $\mathcal{P}$, since $\mathcal{T}_D$ is a rooted tree. If $\mathcal{P}$ is t-local, then whether $\mathcal{T}_D$ has $\mathcal{P}$ depends only on $\mathcal{T}_D|_t$.

Lemma 4 immediately implies the following result, of which Corollary 5 is a special case.

Corollary 6. *Let $\mathcal{P}$ be a local property of rooted graphs, let $D \in \mathcal{D}$, suppose that $\mathbf{d}_n \to D$, and let v be a vertex of $G^*_{\mathbf{d}_n}$ chosen uniformly at random. Then*

$$\mathbb{P}\big((G^*_{\mathbf{d}_n}, v) \in \mathcal{P}\big) \to \mathbb{P}(\mathcal{T}_D \in \mathcal{P})$$

as $n \to \infty$. Equivalently, $\mathbb{E}(N_{\mathcal{P}}(G^*_{\mathbf{d}_n})) = \mathbb{P}(\mathcal{T}_D \in \mathcal{P})|\mathbf{d}_n| + o(|\mathbf{d}_n|)$. □

When we come to concentration, it will be convenient to work with a restatement of this last corollary.

Corollary 7. *Let $\mathcal{P}$ be a local property of rooted graphs, and let $D \in \mathcal{D}$ and $\varepsilon > 0$ be given. Then there exists $\delta > 0$ such that if $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$ then*

$$\big|\mathbb{E}(N_{\mathcal{P}}(G^*_{\mathbf{d}})) - \mathbb{P}(\mathcal{T}_D \in \mathcal{P})n\big| \leqslant \varepsilon n, \tag{11}$$

where $n = |\mathbf{d}|$.

Proof. Suppose not. Then for each n there is a degree sequence $\mathbf{d}_n$ with $d_{\mathrm{conf}}(\mathbf{d}_n, D) \leqslant 1/n$ for which (11) fails. Recalling (5), applying Corollary 6 to $(\mathbf{d}_n)_{n=1}^{\infty}$ gives a contradiction. □

The key property to which we shall apply this result is the property $\mathcal{P}_k$ that the component of the root contains exactly k vertices. Note that in this case

$$N_{\mathcal{P}_k}(G) = N_k(G) \qquad \text{and} \qquad \mathbb{P}(\mathcal{T}_D \in \mathcal{P}_k) = \rho_k(D). \tag{12}$$

We can easily use the second moment method (exploring from two random vertices v and w) to prove that $N_{\mathcal{P}}(G^*_{\mathbf{d}})$ is concentrated in the sense that $N_{\mathcal{P}}(G^*_{\mathbf{d}_n})/n$ converges in probability when $\mathbf{d}_n \to D$ with $|\mathbf{d}_n| = n$. Instead we use the Hoeffding–Azuma inequality to prove a stronger result.

Two configurations (pairings) π_1 and π_2 are *related by a switching* if π_2 can be obtained from π_1 by deleting two pairs $\{a,b\}$ and $\{c,d\}$ and inserting the pairs $\{a,c\}$ and $\{b,d\}$. A function f defined on pairings of some fixed set is *C-Lipschitz* if $|f(\pi_1) - f(\pi_2)| \leqslant C$ whenever π_1 and π_2 are related by a switching. We shall use the following standard simple lemma.

Lemma 8. *Let S be a set with size $2m$, and let f be a C-Lipschitz function of pairings of S. If π is chosen uniformly at random from all pairings of S, then for any $t \geqslant 0$ we have*

$$\mathbb{P}\Big(\big|f(\pi) - \mathbb{E}(f(\pi))\big| \geqslant t\Big) \leqslant 2\exp(-t^2/(4C^2 m)).$$

Proof. Let $S = \{s_1, \ldots, s_{2m}\}$. Let us condition on the partners of $s_1, \ldots, s_i$, writing Ω' for the set of all pairings consistent with the information revealed so far. Now consider s_{i+1}. It may be that its partner is determined, since it is paired to one of $s_1, \ldots, s_i$. If not, for any possible partner b let Ω'_b be the subset of Ω' containing all pairings in which s_{i+1} is paired with b. For distinct possible partners b and c, there is a bijection between Ω'_b and Ω'_c in which each $\pi_1 \in \Omega'_b$ is related to its image π_2 by a switching: we simply switch the pairs $\{s_{i+1}, b\}$ and $\{c,d\}$ for $\{s_{i+1}, c\}$ and $\{b,d\}$, where d is the partner of c in π_1 (and hence of b in π_2).

Write $\mathcal{F}_i$ for the (finite) sigma-field generated by the random variables listing the partners of $s_1, \ldots, s_i$. The bijection just given and the Lipschitz property of f easily imply that $\mathbb{E}(f(\pi) \mid \mathcal{F}_{i+1})$ is always within C of $\mathbb{E}(f(\pi) \mid \mathcal{F}_i)$. Thus the sequence $(X_i)_{i=0}^{2m}$ with $X_i = \mathbb{E}(f(\pi) \mid \mathcal{F}_i)$ is a martingale with differences bounded by C. The result now follows from the Hoeffding–Azuma inequality, noting that $X_0 = \mathbb{E}(f(\pi))$ and $X_{2m} = f(\pi)$. □

Since $N_k(G)$ changes by at most $2k$ when an edge is added to or deleted from a multigraph G, and a switching corresponds to deleting two edges and adding two edges, $N_k(G^*_{\mathbf{d}})$ is $8k$-Lipschitz as a function of the pairing used to generate $G^*_{\mathbf{d}}$. (In fact, it is $4k$-Lipschitz.) Thus Lemma 8 implies concentration of $N_k(G^*_{\mathbf{d}}) = N_{\mathcal{P}_k}(G^*_{\mathbf{d}})$. Later we shall consider more general properties than $\mathcal{P}_k$, and then we must work harder to obtain concentration results – in general for a local property $\mathcal{P}$, there is no constant $C = C(\mathcal{P})$ such that $N_{\mathcal{P}}(G)$ is C-Lipschitz. So we need to modify our properties slightly, to 'avoid high-degree vertices'.

For $\Delta \geqslant 2$ and $t \geqslant 0$, let $\mathcal{M}_{\Delta,t}$ be the property that every vertex within graph distance t of the root has degree at most Δ. Note that $\mathcal{M}_{\Delta,t}$ is $(t+1)$-local.

Lemma 9. *Let $\mathcal{P}$ be a t-local property, and let $\mathcal{Q} = \mathcal{P} \cap \mathcal{M}_{\Delta,t}$. Then the number $N_{\mathcal{Q}}(G)$ of vertices of a multigraph G with property $\mathcal{Q}$ changes by at*

most $4\Delta^t$ *if a single edge is added to or deleted from* G*. Furthermore,* $N_{\mathcal{Q}}(G)$ *is* $16\Delta^t$*-Lipschitz.*

Proof. The effect of a switching on the corresponding configuration multigraph is to delete two edges and then add two edges (perhaps between the same vertices). Thus it suffices to prove the first statement.

Let v be a vertex of G such that one of (G, v) and $(G+e, v)$ has property $\mathcal{Q}$ but the other does not. Note that since $\mathcal{M}_{\Delta,t}$ is monotone decreasing, $(G, v) \in \mathcal{M}_{\Delta,t}$. If $e = xy$, then the graph distance from v to $\{x, y\}$ is the same in G and in $G + e$. Clearly, this distance is at most t; otherwise the presence of e would not affect the property $\mathcal{Q}$. Hence, in G, at least one endpoint of e is within distance t of v, so v is joined to an endpoint of e by a path in G of length at most t in which (since $(G, v) \in \mathcal{M}_{\Delta,t}$) each vertex has degree at most Δ. Each endpoint of e is the end of at most $(1 + \Delta + \cdots + \Delta^t) \leqslant 2\Delta^t$ such paths, so there can be at most $4\Delta^t$ vertices v with the claimed property. $\square$

The next lemma shows that provided we choose Δ large enough, there is no harm in considering only vertices whose local neighbourhoods contain only vertices with degree at most Δ.

Lemma 10. *Let* $D \in \mathcal{D}$*,* $t \geqslant 0$ *and* $\varepsilon > 0$ *be given. Then there exist* $\delta > 0$ *and an integer* Δ *such that*

$$\mathbb{P}(\mathcal{T}_D \textit{ has } \mathcal{M}_{\Delta,t}) \geqslant 1 - \varepsilon/10$$

and

$$\mathbb{P}\big(N_{\mathcal{M}_{\Delta,t}}(G^*_{\mathbf{d}}) \leqslant n - \varepsilon n/2\big) \leqslant e^{-\delta n} \tag{13}$$

whenever $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$*, where* $n = |\mathbf{d}|$*.*

Thus for any given t and ε there is a Δ such that with very high probability, for $d_{\mathrm{conf}}(\mathbf{d}, D)$ small enough, at most $\varepsilon n/2$ vertices of $G^*_{\mathbf{d}}$ are within distance t of a vertex with degree larger than Δ.

Proof. The first statement is immediate from the fact that the random variable M giving the maximum degree of any vertex of $\mathcal{T}_D$ within distance t of the root is always finite, so there is some Δ such that $\mathbb{P}(M > \Delta) < \varepsilon/10$. Corollary 7 implies that, if δ is small enough, then $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$ implies that $N = N_{\mathcal{M}_{\Delta,t}}(G^*_{\mathbf{d}})$ has expectation within $\varepsilon n/10$ of $n\mathbb{P}(\mathcal{T}_D \in \mathcal{M}_{\Delta,t})$, so $\mathbb{E}(N) \geqslant n - \varepsilon n/5$. By Lemma 9, applied with $\mathcal{P}$ the 'trivial' t-local property that always holds, as a function of the pairing used to generate $G^*_{\mathbf{d}}$, the quantity N is C-Lipschitz for some C. Now (13) follows by Lemma 8. $\square$

We are now in a position to establish concentration of the number of vertices whose neighbourhoods have some local property.

Theorem 11. *Let* $\mathcal{P}$ *be a local property of rooted graphs, let* $D \in \mathcal{D}$ *and let* $\varepsilon > 0$*. There is some* $\delta > 0$ *such that if* $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$ *then*

$$\mathbb{P}\Big(\big|N_{\mathcal{P}}(G^*_{\mathbf{d}}) - n\mathbb{P}(\mathcal{T}_D \in \mathcal{P})\big| \geqslant \varepsilon n\Big) \leqslant e^{-\delta n}, \tag{14}$$

where $n = |\mathbf{d}|$.

Proof. Let $D \in \mathcal{D}$, $\varepsilon > 0$ and a t-local property $\mathcal{P}$ be given, and let Δ be as in Lemma 10. Let us say that an event holds *with very high probability* or *wvhp* if for some constant $\delta > 0$ it has probability at least $1 - e^{-\delta n}$ whenever $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$. So in particular, Lemma 10 tells us that wvhp all but at most $\varepsilon n/2$ vertices of $G = G^*_{\mathbf{d}}$ have property $\mathcal{M} = \mathcal{M}_{\Delta,t}$.

Let $N = N_{\mathcal{P}}(G)$ be the number of vertices with property $\mathcal{P}$, let $B = n - N_{\mathcal{M}}(G)$ be the number of 'bad' vertices, i.e, ones not having property $\mathcal{M}$, and let $N' = N_{\mathcal{P}\cap\mathcal{M}}$ be the number of 'good' vertices with property $\mathcal{P}$. Then, wvhp,

$$|N - N'| \leqslant |B| \leqslant \varepsilon n/2.$$

By the first part of Lemma 10, we have

$$|\mathbb{P}(\mathcal{T}_D \in \mathcal{P}) - \mathbb{P}(\mathcal{T}_D \in \mathcal{P} \cap \mathcal{M})| \leqslant \mathbb{P}(\mathcal{T}_D \notin \mathcal{M}) \leqslant \varepsilon/10.$$

By Lemma 9, N' is C-switching Lipschitz for some constant C, so by Corollary 7 and Lemma 8, we have that wvhp

$$|N' - n\mathbb{P}(\mathcal{T}_D \in \mathcal{P} \cap \mathcal{M})| \leqslant \varepsilon n/10,$$

say. The last three displayed equations and the triangle inequality establish (14). □

Corollary 12. *Let* $D \in \mathcal{D}$*, and let* $k \geqslant 1$ *and* $\varepsilon > 0$ *be given. Then there exists* $\delta > 0$ *such that if* $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$ *then*

$$\mathbb{P}\Big(\big|N_k - \rho_k n\big| \geqslant \varepsilon n\Big) \leqslant e^{-\delta n} \tag{15}$$

where $n = |\mathbf{d}|$*,* $N_k = N_k(G^*_{\mathbf{d}})$ *and* $\rho_k = \mathbb{P}(|\mathcal{T}_D| = k)$*.*

Proof. Recall (12) and apply Theorem 11 to the property $\mathcal{P}_k$. □

This corollary proves the first statement (8) of Theorem 2, and hence the corresponding statement in Theorem 1. One can obtain an explicit constant in the exponential error probability in (15) by using that N_k is $4k$-Lipschitz, but there does not seem to be much point.

To conclude this section, we note that, as usual, summing over $k' < k$ and subtracting from n, bounds on N_k with k fixed give bounds on $N_{\geqslant k}$ as well, where $N_{\geqslant k}(G)$ denotes the number of vertices of a graph G in components of order at least k.

Lemma 13. *Let* $D \in \mathcal{D}$*,* $\varepsilon > 0$ *and* K *be given. There exist* $k \geqslant K$ *and* $\delta > 0$ *such that if* $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$ *then*

$$\mathbb{P}\Big(\big|N_{\geqslant k} - \rho(D)n\big| \geqslant \varepsilon n\Big) \leqslant e^{-\delta n}, \tag{16}$$

where $n = |\mathbf{d}|$*,* $N_{\geqslant k} = N_{\geqslant k}(G^*_{\mathbf{d}})$ *and* $\rho(D) = \mathbb{P}(\mathcal{T}_D$ *is infinite*$)$*.*

Proof. Since $\sum_k \rho_k(D) = \mathbb{P}(|\mathcal{T}_D| < \infty) = 1 - \rho(D)$, there is some $k \geqslant K$ such that $\sum_{k'=1}^{k-1} \rho_{k'}$ is within $\varepsilon/2$ of $1 - \rho(D)$. The result follows by applying Lemma 12 for each $k' \leqslant k-1$, with $\varepsilon/(2k)$ in place of ε. □

As usual, the result for k fixed extends to the case when $k \to \infty$ slowly, showing, roughly speaking, that the probability that the branching process $\mathcal{T}_D$ is infinite gives the asymptotic proportion of vertices in 'large' components.

3 The survival probability $\rho(D)$

In this brief section we discuss the behaviour of the survival probability $\rho(D)$ of the branching process $\mathcal{T}_D$. The result below is needed in the next section, but also helps to interpret Theorems 1 and 2.

Recall that from generation 1 onwards, $\mathcal{T}_D$ behaves like the Galton–Watson branching process $\mathcal{T}_D^1$ with offspring distribution Z defined by (9), and that $\mathcal{T}_D$ simply consists of a random number N of copies of $\mathcal{T}_D^1$, with N having the distribution D.

Theorem 14. *Let D be any distribution on the non-negative integers with $\mathbb{P}(D \geqslant 3) > 0$ and $\mathbb{E}(D) < \infty$. Then $\rho(D) > 0$ if and only if $\mathbb{E}(D(D-2)) > 0$. Furthermore, writing x_+ for the largest solution in $[0,1]$ to*

$$x = 1 - \sum_{i=1}^{\infty} \frac{i r_i}{\mathbb{E}(D)} (1-x)^{i-1}, \tag{17}$$

where $r_i = \mathbb{P}(D = i)$, we have

$$\rho(D) = 1 - \sum_{i=0}^{\infty} r_i (1-x_+)^i. \tag{18}$$

Finally, suppose that $D_1, D_2, \ldots$ are distributions on the non-negative integers such that $D_n \to D$ in distribution and $\mathbb{E}(D_n) \to \mathbb{E}(D)$. Then $\rho(D_n) \to \rho(D)$ as $n \to \infty$.

Proof. Standard results on Galton–Watson processes tell us that the survival probability of $\mathcal{T}_D^1$ is equal to x_+, the largest solution in $[0,1]$ to (17). Furthermore, since $\mathbb{P}(D \geqslant 3) > 0$ rules out the trivial case $\mathbb{P}(Z = 1) = 1$, we have $x_+ > 0$ if and only if $\mathbb{E}(Z) > 1$. Conditioning on the number N of children of the root of $\mathcal{T}_D$ gives (18) as an immediate consequence, and shows that $\rho(D) > 0$ if and only if $x_+ > 0$, i.e., if and only if $\mathbb{E}(Z) > 1$. A little manipulation shows that this is equivalent to $\sum_i i(i-2) r_i > 0$.

For the last part, define Z_n from D_n in analogy with (9). Since $\mathbb{P}(D_n = i) \to r_i$ and $\mathbb{E}(D_n) \to \mathbb{E}(D)$, we have $\mathbb{P}(Z_n = i) \to \mathbb{P}(Z = i)$. Standard branching process results then imply that the survival probability of $\mathcal{T}_{D_n}^1$ converges to that of $\mathcal{T}_D^1$. Using (18), it follows easily that $\rho(D_n) \to \rho(D)$. □

Remark 15. The formulae above coincide (as they must) with those given by Molloy and Reed [12] – one can check that x_+ is equal to $1 - \sqrt{1 - 2\alpha_D/K}$ in their notation. They did not use the branching process interpretation, however. In the notation of Janson and Luczak [9], x_+ is $1 - \xi$, and $\rho(D)$ is $1 - g(\xi)$.

4 Colouring and sprinkling

Our next task is to use 'sprinkling' to show that whp almost all vertices in 'large' components are in a single 'giant' component. In the original context of the random graphs $G(n, p)$ and $G(n, m)$, sprinkling is very simple to implement – first include each edge independently with probability p_1, then 'sprinkle' in extra edges by including each edge not already present independently with probability p_2, where $p_1 + p_2 - p_1 p_1 = p$. In the context of the configuration model, there is no very simple analogue of this. Instead, we will 'thin' the random graph $G^*_{\mathbf{d}}$, and then put back the deleted edges.

Given $0 < p < 1$, let $G' = G^*_{\mathbf{d}}[p]$ denote the random subgraph of $G = G^*_{\mathbf{d}}$ obtained by retaining each edge independently with probability p, and let G'' be the multigraph formed by the deleted edges, so $V(G'') = V(G') = V(G)$ and $E(G)$ is the disjoint union of $E(G')$ and $E(G'')$. Let $\mathbf{d}'$ be the (random, of course) degree sequence of G', and $\mathbf{d}''$ that of G'', so $d'_i + d''_i = d_i$ for each vertex $i \in V(G)$. The following simple observation is a key ingredient of the sprinkling argument.

Lemma 16. *For any* $\mathbf{d}$ *and any* $0 < p < 1$*, the random graphs* G' *and* G'' *are conditionally independent given* $\mathbf{d}'$*, having the distributions of* $G^*_{\mathbf{d}'}$ *and* $G^*_{\mathbf{d}''}$ *respectively.*

Proof. This is essentially immediate from the definition of the configuration model: recall that G is defined from a pairing π of a set of $2m(\mathbf{d})$ stubs. Given this pairing, colour each pair red with probability p and blue otherwise, independently of the others. Then we may take G' to be given by the red pairs and G'' by the blue pairs. Clearly, given the set of stubs in red pairs (which determines $\mathbf{d}'$ and thus $\mathbf{d}''$), the pairing of these red stubs is uniformly random, and similarly for the blue stubs. □

Our next aim is to extend the coupling result Theorem 11 to the pair (G', G''). First we need some definitions. We shall work with 2-coloured multigraphs (rather than coloured pairings as above). Given a degree sequence $\mathbf{d}$ and $0 < p < 1$, let $G^*_{\mathbf{d}}\{p\}$ denote the random *coloured* graph obtained by constructing $G^*_{\mathbf{d}}$ and then colouring the edges independently, each red with probability p and blue otherwise. Thus $G' = G^*_{\mathbf{d}}[p]$ may be viewed as the red subgraph of $G^*_{\mathbf{d}}\{p\}$. Similarly, let $\mathcal{T}_D\{p\}$ be the random coloured rooted tree obtained from $\mathcal{T}_D$ by colouring each edge red with probability p and blue otherwise, independently of the others.

Given a probability distribution D on the non-negative integers, and $0 < p < 1$, let D_p be the *p-thinned* version of D, which may be defined by taking a random

set X of size D and selecting elements of X independently with probability p. Then D_p is the (overall) distribution of the number of selected elements. It is a simple exercise in basic probability to check that the operations of (i) p-thinning and (ii) size-biasing and then subtracting 1 commute. A simple consequence of this is that the component of the red subgraph of $\mathcal{T}_D\{p\}$ containing the root has the same distribution as $\mathcal{T}_{D_p}$.

The next result concerns 'local properties of coloured rooted graphs', which are defined in the obvious way.

Theorem 17. *Let $\mathcal{P}$ be a local property of coloured rooted graphs, let $D \in \mathcal{D}$, let $\varepsilon > 0$ and let $0 < p < 1$. There is some $\delta > 0$ such that if $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$ then*

$$\mathbb{P}\Big(\big|N_{\mathcal{P}}(G^*_{\mathbf{d}}\{p\}) - n\mathbb{P}(\mathcal{T}_D\{p\} \in \mathcal{P})\big| \geqslant \varepsilon n\Big) \leqslant e^{-\delta n}, \tag{19}$$

where $n = |\mathbf{d}|$. Furthermore, if $\mathcal{Q}$ is a local property of rooted graphs, then there is some $\delta > 0$ such that if $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$ then

$$\mathbb{P}\Big(\big|N_{\mathcal{Q}}(G^*_{\mathbf{d}}[p]) - n\mathbb{P}(\mathcal{T}_{D_p} \in \mathcal{Q})\big| \geqslant \varepsilon n\Big) \leqslant e^{-\delta n}. \tag{20}$$

Proof. From the remarks above, it suffices to prove the first statement, (19). Then (20) may be deduced by applying (19) to the local property $\mathcal{P}$ that the component of the red graph containing the root has property $\mathcal{Q}$. We give only an outline proof of (19), since the argument is a simple modification of that of Theorem 11.

Firstly, the coloured analogue of Lemma 4 follows from Lemma 4: when the coupling as uncoloured graphs succeeds, we may apply the same (random) colouring to $\Gamma_{\leqslant t}(v)$ as to $\mathcal{T}_D|_t$. Arguing as before, we deduce the coloured analogue of Corollary 7. Now $N_{\mathcal{P}}(G^*_{\mathbf{d}}\{p\})$ depends not only on the configuration, but also on the colouring. However, passing to a property $\mathcal{Q} = \mathcal{P} \cap \mathcal{M}_{\Delta,t}$ as in the proof of Theorem 17, by a variant of Lemma 9 we see that $N_{\mathcal{Q}}$ changes by at most a constant (a) under a switching and (b) under changing the colour of a single edge. Now we can apply the Hoeffding–Azuma inequality to a martingale with $2m$ steps for the switchings and m for the colour choices, where $m = m(\mathbf{d})$ is the number of edges of $G^*_{\mathbf{d}}$, to deduce concentration of $N_{\mathcal{P}}(G^*_{\mathbf{d}}\{p\})$ and complete the proof. □

Recall that $n_i = n_i(\mathbf{d})$ is the number of vertices with degree i in $G = G^*_{\mathbf{d}}$. Let n'_i be the number of vertices with degree i in the random subgraph $G' = G[p]$ defined earlier. Also, for $0 \leqslant i \leqslant j$, let n_{ij} be the number of vertices with degree i in G' and degree j in G. Thus $n'_i = \sum_{j \geqslant i} n_{ij}$.

Writing $r_i = \mathbb{P}(D = i)$ as usual, for $0 \leqslant i \leqslant j$ let

$$r_{ij} = r_j \binom{j}{i} p^i (1-p)^{j-i}, \tag{21}$$

and let

$$r'_i = \sum_{j \geqslant i} r_{ij} = \mathbb{P}(D_p = i). \tag{22}$$

Informally, these expressions give the expected proportions of vertices of G' having degree i (for r'_i) and having degree i in G' and degree j in G, ignoring the effect of loops.

Lemma 18. *Let $D \in \mathcal{D}$ and $0 < p < 1$ be fixed. Given $0 \leqslant i \leqslant j$ and $\varepsilon > 0$ there exists $\delta > 0$ such that if $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$ then*

$$\mathbb{P}\big(|n_{ij} - r_{ij}n| \geqslant \varepsilon n\big) \leqslant e^{-\delta n}$$

and

$$\mathbb{P}\big(|n'_i - r'_i n| \geqslant \varepsilon n\big) \leqslant e^{-\delta n},$$

where $n = |\mathbf{d}|$.

Proof. Apply Theorem 17 to the 1-local coloured rooted graph properties 'the root is incident with j edges in total of which i are red' for the first statement, and 'the root is incident with i' red edges' for the second. □

Recall that D_p is the p-thinned version of the probability distribution D, which may be defined by (22). Lemma 18 has the following consequence.

Corollary 19. *Given $D \in \mathcal{D}$, $0 < p < 1$ and $\varepsilon > 0$ there exists $\delta > 0$ such that, if $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$, then*

$$\mathbb{P}\Big(d_{\mathrm{conf}}(\mathbf{d}', D_p) \geqslant \varepsilon\Big) \leqslant e^{-\delta n},$$

where $\mathbf{d}'$ is the degree sequence of the random subgraph $G[p]$ of $G = G^*_{\mathbf{d}}$ and $n = |\mathbf{d}| = |\mathbf{d}'|$ is the number of vertices.

Proof. Since $\mathbb{E}(D) < \infty$ there is a constant C such that $\sum_{i \geqslant C} i r_i < \varepsilon/8$. If δ is small enough, then $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$ implies $\sum_{i \geqslant C} i n_i(\mathbf{d}) < \varepsilon n/4$. Since D stochastically dominates D_p, and the degree of a vertex in our random subgraph G' is at most its degree in G, the corresponding bounds for D_p and $n'_i = n_i(\mathbf{d}'_n)$ follow. From the definition (3), (4) of d_{conf} it thus suffices to prove that for each fixed $i < C$ we have $|n'_i - r'_i n| \leqslant \varepsilon/(2C^2)$ with sufficiently high probability; this follows from Lemma 18. □

The next trivial lemma will be applied to the sprinkled edges.

Lemma 20. *Let A and B be disjoint sets of stubs in the configuration model associated to $G^*_{\mathbf{d}}$. Then the probability that no stubs in A are paired to stubs in B is at most $\exp(-|A||B|/(8m))$, where $m = m(\mathbf{d})$.*

Proof. Assume without loss of generality that $|A| \leqslant |B|$. Perform a sequence of $\lceil |A|/2 \rceil$ experiments, each consisting of choosing an as-yet-unpaired stub in A and revealing its partner. In the ith experiment, there are at least $|B| - (\lceil |A|/2 \rceil - 1) \geqslant |B| - |A|/2 \geqslant |B|/2$ unpaired stubs in B, so the probability of finding the partner in B is at least $(|B|/2)/(2m + 1 - 2i) \geqslant |B|/(4m)$. Hence the probability that no partner in B is found is at most $(1 - |B|/(4m))^{|A|/2} \leqslant \exp(-|A||B|/(8m))$. □

We are finally ready to prove the multigraph case of Theorem 2, where $G_{\mathbf{d}}$ is replaced by $G^*_{\mathbf{d}}$.

Proof of Theorem 2 for $G^*_{\mathbf{d}}$. Let $L_i = L_i(G^*_{\mathbf{d}})$ be the number of vertices in the ith largest component of $G^*_{\mathbf{d}}$.

Fix $D \in \mathcal{D}$ and $\varepsilon > 0$. By Lemma 13 there are constants k and $\delta > 0$ such that if $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$, then $\mathbb{P}\big(N_{\geqslant k}(G) \geqslant (\rho(D) + \varepsilon/8)n\big) \leqslant e^{-\delta n}$. Since $L_1 + L_2 \leqslant N_{\geqslant k} + 2k$, if n is large enough (which we can ensure by taking δ small enough) it follows that

$$\mathbb{P}\Big(L_1 + L_2 \leqslant (\rho(D) + \varepsilon/4)n\Big) \geqslant 1 - e^{-\delta n}.$$

To complete the proof, it suffices to show that, reducing δ if necessary, if $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$ then

$$\mathbb{P}\Big(L_1 \geqslant (\rho(D) - 3\varepsilon/4)n\Big) \geqslant 1 - e^{-\delta n}. \tag{23}$$

As $p \to 1$, the probability distribution D_p defined above converges to D, both in distribution and (since $\mathbb{E}(D_p) \leqslant \mathbb{E}(D) < \infty$) in expectation. Hence, Theorem 14 tells us that $\rho(D_p) \to \rho(D)$ as $p \to 1$. (This is the only place in the argument where $\mathbb{P}(D \geqslant 3) > 0$ is used.) In particular, there is some $p < 1$ such that

$$\rho(D_p) \geqslant \rho(D) - \varepsilon/8.$$

Let us fix such a p for the rest of the proof. Also, fix an integer $t \geqslant 1$ such that

$$p^t \leqslant \varepsilon/20,$$

set

$$K = 1 + \Delta + \cdots + \Delta^{t-1} + 1,$$

and let

$$\alpha = \frac{\varepsilon}{40\Delta^t} \qquad \text{and} \qquad \gamma = \frac{\alpha^2}{8\mathbb{E}(D)}. \tag{24}$$

We shall study the coloured random graph $G^*_{\mathbf{d}}\{p\}$ defined earlier, obtained from $G^*_{\mathbf{d}}$ by colouring each edge red with probability p and blue otherwise, independently of the others. As before, we write $G' = G^*_{\mathbf{d}}[p]$ for the red subgraph and G'' for the blue subgraph, and $\mathbf{d}'$ and $\mathbf{d}''$ for the degree sequences of G' and G''. Recall that, by Lemma 16, given $\mathbf{d}'$, we can view G' and G'' as independent configuration multigraphs.

Applying Lemma 13 to G', we find that there exist $k \geqslant \max\{K, 2/\gamma\}$ and $\delta_1 > 0$ such that, writing L for the set of vertices in components of G' with at least k vertices, we have

$$\mathbb{P}\Big(|L| \geqslant (\rho(D) - \varepsilon/4)n \;\Big|\; \mathbf{d}'\Big) \geqslant \mathbb{P}\Big(|L| \geqslant (\rho(D_p) - \varepsilon/8)n \;\Big|\; \mathbf{d}'\Big) \geqslant 1 - e^{-\delta_1 n}$$

whenever $d_{\mathrm{conf}}(\mathbf{d}', D_p) < \delta_1$. By Corollary 19 there is a $\delta_2 > 0$ such that if $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta_2$, then

$$\mathbb{P}\Big(d_{\mathrm{conf}}(\mathbf{d}', D_p) \geqslant \delta_1\Big) \leqslant e^{-\delta_2 n}.$$

Hence, reducing δ if necessary, it follows that if $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$ then

$$\mathbb{P}\Big(|L| \geqslant (\rho(D) - \varepsilon/4)n\Big) \geqslant 1 - e^{-\delta_1 n} - e^{-\delta_2 n} \geqslant 1 - e^{-\delta n}. \tag{25}$$

Note that in the argument above we could have sidestepped Corollary 19, using a coloured version of Theorem 13 and considering the coloured property 'the red component of the root has size at least k'. However, the approach above seems more intuitive and we shall use Corollary 19 in Section 6.

Let us call a vertex $v \in V(G) = V(G')$ *usable* if it is incident with a blue edge, i.e., an edge of G''. (These edges will be our 'sprinkled' edges.) Note that knowing $\mathbf{d}'$ determines whether v is usable: we don't know which edges are present in G'', but we do know its degree sequence. Our next aim is to find 'enough' usable vertices in L, for which we need some further definitions.

By the *radius* $r(G)$ of a (locally finite) rooted graph G we mean the maximum distance of any vertex from the root, considering only vertices in the component C containing the root. Thus $r(G)$ is infinite if and only if C is infinite.

Given a coloured rooted graph G, we write $R(G)$ and $B(G)$ for the red and blue subgraphs of G, respectively. Let $\mathcal{G}_t$ be the property of coloured rooted graphs G that either

(i) $r(R(G)) < t$ or

(ii) some vertex of G within distance t of the root is incident with an edge of $B(G)$.

Note that, considering the shortest path to a blue edge, (ii) is equivalent to (ii') some vertex of $R(G)$ within distance t (in $R(G)$) of the root is incident with an edge of $B(G)$. The property $\mathcal{G}_t$ is clearly $(t+1)$-local.

Consider the case where $G = \mathcal{T}_D\{p\}$ is a coloured rooted tree. Conditioning first on the graph structure, if $r(G) < t$ then (i) will certainly hold. Otherwise, there are at least t edges of G within distance t of the root, and if any one is blue (ii) holds. Thus

$$\mathbb{P}(\mathcal{T}_D\{p\} \in \mathcal{G}_t) \geqslant 1 - p^t \geqslant 1 - \varepsilon/20.$$

By Lemma 10 (with $\varepsilon/2$ in place of ε), there is some Δ such that $\mathbb{P}(\mathcal{T}_{D_p} \in \mathcal{M}_{\Delta,t}) \geqslant 1 - \varepsilon/20$. Let $\mathcal{H}$ be the property

$$\mathcal{H} = \{R(G) \in \mathcal{M}_{\Delta,t} \text{ and } G \in \mathcal{G}_t\},$$

noting that

$$\mathbb{P}(\mathcal{T}_D\{p\} \in \mathcal{H}) \geqslant 1 - \varepsilon/10. \tag{26}$$

We call a vertex v of our coloured configuration model $G = G^*_{\mathbf{d}}\{p\}$ *helpful* if $(G, v) \in \mathcal{H}$, i.e., if G rooted at v has property $\mathcal{H}$. Let H denote the set of

helpful vertices. From (26) and Theorem 17, if δ is chosen small enough, then if $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$ we have

$$\mathbb{P}\big(|H| \leqslant n - \varepsilon n/5\big) \leqslant e^{-\delta n}. \tag{27}$$

Since, as noted above, knowing $\mathbf{d}'$ determines which vertices are usable (incident with edges of G''), it is easy to check from the definition of $\mathcal{H}$ that knowing $\mathbf{d}$ (which is given), $\mathbf{d}'$ and G' determines which vertices of G are helpful.

From now on we condition on $\mathbf{d}'$ and G', assuming that

$$|L| \geqslant (\rho(D) - \varepsilon/4)n \qquad \text{and} \qquad |H| \geqslant n - \varepsilon n/5. \tag{28}$$

This makes sense since L (the set of vertices in components of G' with order at least k) and H are determined by $\mathbf{d}'$ and G', and (25) and (27) imply that the event (28) has probability at least $1 - e^{-\delta n}$.

Suppose that $v \in L \cap H$. Then, since v is helpful, every vertex in the t-neighbourhood $\Gamma^{G'}_{\leqslant t}(v)$ of v in G' has degree at most Δ. Furthermore, from the definition of $\mathcal{G}_t$ (recalling (ii') above), either (a) the radius of G' rooted at v is at most $t-1$, or (b) $\Gamma^{G'}_{\leqslant t}(v)$ meets an edge of G'', i.e., contains a usable vertex. In case (a), it follows that the component of v in G' has at most $1 + \Delta + \cdots + \Delta^{t-1} < K$ vertices, contradicting $v \in L$. Thus case (b) holds and there is a path $P_v = v_0 v_1 \cdots v_r$ in G' of length at most t where $v_0 = v$, each v_i has degree at most Δ in G', and v_r is usable.

At this point we are finally ready to apply the sprinkling strategy of Erdős and Rényi [4]. Let us call a partition (X, Y) of L a *potentially bad cut* if $|X|$, $|Y| \geqslant \varepsilon n/4$ and there are no edges of G' joining X to Y. We call (X, Y) a *bad cut* if, in addition, no edge of G'' joins X to Y. Since each component of G' in L must lie either entirely in X or entirely in Y, there are at most

$$2^{|L|/k} \leqslant 2^{n/k} \leqslant e^{n/k} \leqslant e^{\gamma n/2} \tag{29}$$

potentially bad cuts, recalling that we chose $k \geqslant 2/\gamma$.

Let (X, Y) be a potentially bad cut, and recall that $|H| \geqslant n - \varepsilon n/5$. Thus X contains at least $\varepsilon n/20$ helpful vertices v. From each there is a path P_v as described above ending at some usable vertex u. Because of the degree conditions, at most $1 + \Delta + \cdots + \Delta^t \leqslant 2\Delta^t$ such paths can end at a given usable vertex. Since P_v is a path in G', and X is a union of components of G', we conclude that X contains at least αn usable vertices, where $\alpha = \varepsilon/(40\Delta^t)$ as in (24). Of course, the same applies to Y.

Recall that we have conditioned on $\mathbf{d}'$ and G', but not on G''. In the configuration model corresponding to G'', each usable vertex has at least one stub, so X and Y each correspond to sets of at least αn stubs. Since (if δ is chosen small enough) G'' has at most $\mathbb{E}(D)n$ edges, by Lemma 20

$$\mathbb{P}\big(G'' \text{ contains no } (X,Y) \text{ edge} \mid \mathbf{d}', G'\big) \leqslant e^{-\frac{\alpha^2 n^2}{8\mathbb{E}(D)n}} = e^{-\gamma n}.$$

From (29) it follows that the expected number of bad cuts (given $\mathbf{d}'$ and G') is at most $e^{-\gamma n/2}$, so the probability that there are any bad cuts is at most $e^{-\gamma n/2}$.

When there are no bad cuts, it is easy to check that $L_1(G) \geqslant |L| - 2\varepsilon n/4 \geqslant (\rho(D) - 3\varepsilon/4)n$, completing the proof of (23) and hence of the multigraph case of Theorem 2. $\square$

5 Simple graphs

As noted in the introduction, Janson and Luczak [9] proved a result that is similar to the multigraph case of Theorem 2: the assumptions are identical, but the error bounds in the conclusions in [9] are much weaker. An advantage of our stronger error bounds is that they allow us to translate the result to random *simple* graphs without further restrictions on the degree sequences. For this we need a simple lemma.

Lemma 21. *Let $D \in \mathcal{D}$. Then for any $\varepsilon > 0$ there exists a $\delta > 0$ such that if $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$ then*

$$\mathbb{P}\big(G^*_{\mathbf{d}} \text{ is simple }\big) \geqslant e^{-\varepsilon n},$$

where $n = |\mathbf{d}|$. Equivalently, if $D \in \mathcal{D}$ and $\mathbf{d}_n \to D$ in the sense that (1) *and* (2) *hold and $|\mathbf{d}_n| \to \infty$, then*

$$\mathbb{P}\big(G^*_{\mathbf{d}_n} \text{ is simple }\big) = e^{-o(|\mathbf{d}_n|)}.$$

In particular, the degree sequences we consider here are (for large n) realizable by simple graphs.

Proof. The equivalence of the two statements follows easily from (5); we prove the first form.

Observe that there are constants K, M and $\alpha > 0$ such that, if δ is chosen small enough, then $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$ ensures that at least αn vertices of $\mathbf{d}$ have degree between 1 and K (inclusive), and $m = m(\mathbf{d}) \leqslant Mn$, where $n = |\mathbf{d}|$ as usual. Indeed, choose any $K \geqslant 1$ such that $\mathbb{P}(D = K) > 0$, let $\delta \leqslant \alpha = \mathbb{P}(D = K)/2$, and take $M = \mathbb{E}(D)/2 + \alpha$, say. These properties and (6)/(7) are all that we need to know about $\mathbf{d}$.

Let $\mathcal{S}$ denote the event that $G^*_{\mathbf{d}}$ is simple, and fix $\varepsilon > 0$. Pick $\eta > 0$ such that $\eta \log(4M/\alpha) \leqslant \varepsilon/2$ and $\eta \leqslant \alpha/2$. By (7) there is a constant C, which we may take to be larger than K, such that if δ is small enough, then at most ηn stubs are attached to vertices of degree at least C. Let us call a vertex *low degree* if its degree is between 1 and K, and *high degree* if its degree is at least C. Let $\mathcal{E}$ be the event that the stubs attached to high degree vertices are paired with stubs attached to *distinct* low degree vertices.

To determine whether $\mathcal{E}$ holds, we test the at most ηn stubs attached to high degree vertices one-by-one. At each stage, there are at least $\alpha n - \eta n \geqslant \alpha n/2$ low-degree vertices none of whose stubs has yet been paired. Since each such vertex has degree at least one, and there are at most $2Mn$ unpaired stubs in total, it follows that

$$\mathbb{P}(\mathcal{E}) \geqslant \Big(\frac{\alpha n}{4Mn}\Big)^{\eta n} \geqslant e^{-\varepsilon n/2}.$$

When $\mathcal{E}$ holds, the graph $G^*_{\mathbf{d}}$ is simple if and only if the graph G_0 formed by the edges not incident with high-degree vertices is simple. But, after revealing the partners of the stubs attached to the high-degree vertices, the conditional distribution of G_0 is given by the configuration model for some degree sequence in which all degrees are at most C, and at least $\alpha n/2 = \Theta(n)$ degrees are positive. The original result of Bollobás [2] thus gives $\mathbb{P}(\mathcal{S} \mid \mathcal{E}) = \Theta(1)$, and the result follows. □

Proof of Theorem 2 for $G_{\mathbf{d}}$. Let $\mathcal{P}$ be any property of graphs. Since the distribution of $G^*_{\mathbf{d}}$ conditioned on the event $\mathcal{S}$ that $G^*_{\mathbf{d}}$ is simple is exactly that of $G_{\mathbf{d}}$, we have

$$\mathbb{P}(G_{\mathbf{d}} \in \mathcal{P}) = \mathbb{P}(G^*_{\mathbf{d}} \in \mathcal{P} \mid G^*_{\mathbf{d}} \in \mathcal{S}) \leqslant \frac{\mathbb{P}(G^*_{\mathbf{d}} \in \mathcal{P})}{\mathbb{P}(G^*_{\mathbf{d}} \in \mathcal{S})}.$$

Fix $D \in \mathcal{D}$. All statements about $G^*_{\mathbf{d}}$ in Theorem 2 are of the form that for some property $\mathcal{P}$, there exist $\gamma, \delta_1 > 0$ such that if $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta_1$, then $\mathbb{P}(G^*_{\mathbf{d}} \in \mathcal{P}) \leqslant e^{-\gamma n}$. (The theorem asserts this with $\delta_1 = \gamma$.) Lemma 21 gives us $\delta_2 > 0$ such that $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta_2$ implies $\mathbb{P}(G^*_{\mathbf{d}} \in \mathcal{S}) \geqslant e^{-\gamma n/2}$. Hence, setting $\delta = \min\{\delta_1, \delta_2, \gamma/2\}$, if $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$ then

$$\mathbb{P}(G_{\mathbf{d}} \in \mathcal{P}) \leqslant e^{-\gamma n}/e^{-\gamma n/2} = e^{-\gamma n/2} \leqslant e^{-\delta n},$$

completing the proof of Theorem 2. □

As noted in the introduction, Theorem 2 implies Theorem 1.

6 Extensions

One of the motivations for studying the size of the largest component in the configuration model $G_{\mathbf{d}}$ is to consider percolation in this random environment: given $0 < p < 1$, when does the random subgraph $G_{\mathbf{d}}[p]$ of $G_{\mathbf{d}}$ obtained by selecting edges independently with probability p contain a giant component? For example, Goerdt [6] showed that when $G_{\mathbf{d}}$ is simply a random d-regular graph, then there is a 'threshold' at $p = 1/(d-1)$, above which a giant component appears. As is by now well known, for results of the present type this question turns out to be no more general than studying $G_{\mathbf{d}}$ directly (i.e., the case $p = 1$), since one can view a random subgraph of the configuration model as another instance of the configuration model. This is discussed in detail by Fountoulakis [5]; for a slightly different approach see Janson [8]. We give the short argument since it is very easy with the ingredients we have to hand. In the next result we state only the most interesting part formally; D_p is the probability distribution appearing in Corollary 19.

Theorem 22. *Let $0 < p < 1$ be fixed. The conclusions of Theorems 1 and 2 hold if $G^*_{\mathbf{d}}$ or $G_{\mathbf{d}}$ is replaced by its random subgraph $G^*_{\mathbf{d}}[p]$ or $G_{\mathbf{d}}[p]$, and $\rho(D)$ and $\rho_k(D)$ are replaced by $\rho(D_p)$ and $\rho_k(D_p)$.*

In particular, given $D \in \mathcal{D}$ with $\mathbb{P}(D \geqslant 3) > 0$, $0 < p < 1$ and $\varepsilon > 0$, there exists $\delta > 0$ such that, if $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$, then

$$\mathbb{P}\Big(\big|L_1(G_{\mathbf{d}}[p]) - \rho(D_p)n\big| \geqslant \varepsilon n\Big) \leqslant e^{-\delta n}$$

and

$$\mathbb{P}\Big(\big|L_1(G^*_{\mathbf{d}}[p]) - \rho(D_p)n\big| \geqslant \varepsilon n\Big) \leqslant e^{-\delta n},$$

where $n = |\mathbf{d}|$.

Proof. For $G^*_{\mathbf{d}}[p]$, this is essentially trivial from Theorem 2 and Corollary 19. Indeed, by Theorem 2 there exists $\delta_1 > 0$ such that if $d_{\mathrm{conf}}(\mathbf{d}_1, D_p) < \delta_1$ then $G^*_{\mathbf{d}_1}$ has the desired property (L_1 close to $\rho(D_p)n$) with probability at least $1 - e^{-\delta_1 n}$. By Corollary 19 there is a δ such if $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$ then $\mathbb{P}(d_{\mathrm{conf}}(\mathbf{d}', D_p) < \delta_1) \geqslant 1 - e^{-\delta n}$, where $\mathbf{d}'$ is the degree sequence of $G^*_{\mathbf{d}}[p]$. The result for $G^*_{\mathbf{d}}[p]$ follows by noting that, conditional on $\mathbf{d}'$, $G^*_{\mathbf{d}}[p]$ has the distribution of $G^*_{\mathbf{d}'}$.

For $G_{\mathbf{d}}[p]$ we argue as in the last part of the previous section: note that conditional on $G^*_{\mathbf{d}}$ being simple, $G^*_{\mathbf{d}}[p]$ has the same distribution as $G_{\mathbf{d}}[p]$. Then use Lemma 21 as before. The key point is that we do not try to condition on $G^*_{\mathbf{d}}[p]$ being simple. □

Remark 23. Theorem 22 implies that there is a 'critical' p_{c} such that $G^*_{\mathbf{d}}[p]$ has a giant component if and only if $p > p_{\mathrm{c}}$. Indeed, $p_{\mathrm{c}} = \inf\{p : \rho(D_p) = 0\}$. From basic branching process results, it is easy to see that $p_{\mathrm{c}} = 1/\mathbb{E}(Z)$, where Z is the distribution defined in (9). Either from this, or from the fact that $\rho(D_p) > 0$ if and only if $\mathbb{E}(D_p(D_p - 2)) > 0$ it is easy to see that

$$p_{\mathrm{c}} = \frac{\mathbb{E}(D)}{\mathbb{E}(D(D-1))}.$$

This is the same formula as given by Fountoulakis [5], for example, who proved a form of Theorem 22, with stronger assumptions on the degree sequences and weaker error bounds.

Remark 24. Taking $|\mathbf{d}_n| = n$ for notational simplicity, in the context of Theorems 1 and 2, the assumption that $\mathbb{E}(D) < \infty$, corresponding to $m(\mathbf{d}_n) = O(n)$, is very natural. Indeed, it is not hard to see that if $m(\mathbf{d}_n)/n \to \infty$, then $G^*_{\mathbf{d}_n}$ will with high probability contain a component with $n - o(n)$ vertices. As soon as we consider percolation on $G^*_{\mathbf{d}_n}$, however, it makes very good sense to allow $m(\mathbf{d}_n)/n \to \infty$ and then study $G^*_{\mathbf{d}_n}[p_n]$ with $p_n \to 0$ as $n \to \infty$. All we shall say here is that in many situations, for appropriate p_n, the (random) degree sequence of $G^*_{\mathbf{d}_n}[p_n]$ will with high probability be such that Theorem 1 applies to it. For example, if all degrees are equal to k_n with $k_n \to \infty$ and $k_n p_n \to \lambda \in \mathbb{R}$, then the degree distribution of $G^*_{\mathbf{d}_n}[p_n]$ will be asymptotically Poisson with mean λ. Hence Theorem 1 can be used to show that the threshold for percolation on $G^*_{\mathbf{d}_n}$ is at $\lambda = 1$, i.e., at $p_n = 1/k_n$.

Throughout the paper we have focussed on the number of vertices in the giant component. What can we say about other properties, such as the number of vertices of given degree, or the total number of edges? Results for these are given (under different conditions) by Janson and Luczak [9], for example. An often neglected benefit of the branching-process viewpoint is that it typically gives results of this type essentially automatically, not just for these properties, but for any local property. (A version of this observation was made in a different context by Bollobás, Janson and Riordan [3, Lemma 11.11]; see also [14, Theorem 2.8].)

We state the following result in a form analogous to Theorem 2; this of course implies a version analogous to Theorem 1.

Theorem 25. *Let $\mathcal{P}$ be a local property of rooted graphs, let $D \in \mathcal{D}$ and let $\varepsilon > 0$. There is some $\delta > 0$ such that if $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$ then the following hold, with $n = |\mathbf{d}|$ and $G = G^*_{\mathbf{d}}$ or $G = G_{\mathbf{d}}$:*

$$\mathbb{P}\Big(\big|N_{\mathcal{P}}(G) - n\mathbb{P}(\mathcal{T}_D \text{ has } \mathcal{P})\big| \geqslant \varepsilon n\Big) \leqslant e^{-\delta n}, \tag{30}$$

and

$$\mathbb{P}\Big(\big|N_{\mathcal{P}}(C_1) - n\mathbb{P}(\mathcal{T}_D \text{ is infinite and has } \mathcal{P})\big| \geqslant \varepsilon n\Big) \leqslant e^{-\delta n}, \tag{31}$$

where C_1 is a component of G of maximal order.

Proof. As usual, in the light of Lemma 21 we need only consider the case $G = G^*_{\mathbf{d}}$. In this case, we have proved (30) already in Theorem 11.

Turning to (31), let $D \in \mathcal{D}$, $\varepsilon > 0$ and a local property $\mathcal{P}$ be given. Let $\mathcal{S}_k$ be the rooted-graph property 'the component of the root contains at least k vertices', and $\mathcal{S}_\infty$ 'the component of the root is infinite'. (We only consider the latter in the context of $\mathcal{T}_D$; all our graphs here are finite.) Then, as $k \to \infty$, the events $\{\mathcal{T}_D \in \mathcal{S}_k\} = \{|\mathcal{T}_D| \geqslant k\}$ decrease to the event $\{\mathcal{T}_D \in \mathcal{S}_\infty\} = \{\mathcal{T}_D \text{ is infinite}\}$. Hence $\mathbb{P}(\mathcal{T}_D \in \mathcal{S}_k) \to \mathbb{P}(\mathcal{T}_D \in \mathcal{S}_\infty)$, and there is a constant K such that for any $k \geqslant K$ we have

$$\mathbb{P}(\mathcal{T}_D \in \mathcal{S}_k \setminus \mathcal{S}_\infty) < \varepsilon/10. \tag{32}$$

As before, let us say that an event holds 'wvhp' if for some $\delta > 0$ it holds with probability at least $1 - e^{-\delta n}$ whenever $d_{\mathrm{conf}}(\mathbf{d}, D) < \delta$. By Lemma 13 there is some $k \geqslant K$ such that wvhp

$$\big|N_{\geqslant k}(G^*_{\mathbf{d}}) - \rho(D)n\big| \leqslant \varepsilon n/10. \tag{33}$$

Let $N = N_{\mathcal{P}}(C_1)$ be the number of vertices we wish to count, i.e., those in the largest component C_1 of $G^*_{\mathbf{d}}$ having property $\mathcal{P}$. Let $N' = N_{\mathcal{P}\cap\mathcal{S}_k}(G^*_{\mathbf{d}})$ count vertices with property $\mathcal{P}$ in components of size at least k. Then N and N' differ by at most $N_{\geqslant k} - L_1$, which, by (33) and Theorem 2, is wvhp at most $\varepsilon n/5$, say. Applying (30) to the local property $\mathcal{P} \cap \mathcal{S}_k$, we deduce that wvhp N is within $\varepsilon n/4$ of $n\mathbb{P}(\mathcal{T}_D \in \mathcal{P} \cap \mathcal{S}_k)$. But by (32) this is within $\varepsilon n/10$ of $n\mathbb{P}(\mathcal{T}_D \in \mathcal{P} \cap \mathcal{S}_\infty)$, establishing (31). $\square$

For simple properties $\mathcal{P}$, it is easy to give explicit formulae for the probability that $\mathcal{T}_D$ is infinite and has property $\mathcal{P}$. For example, if $\mathcal{P} = \mathcal{P}_d$ is the property that the root has degree d, then defining x_+ as in Section 3, the proof of Theorem 14 shows easily that

$$\mathbb{P}(\mathcal{T}_D \text{ is infinite and has } \mathcal{P}_d) = r_d(1 - (1 - x_+)^d).$$

This gives an asymptotic formula for the number of degree-d vertices in the giant component C_1 that coincides with that of Janson and Luczak [9].

Rather than counting vertices with some local property, what happens if we want to sum some 'local function' $f(G, v)$ over vertices $v \in C_1$? Can we show that

$$n^{-1} \sum_{v \in C_1} f(C_1, v) \xrightarrow{\mathrm{p}} \mathbb{E}(f(\mathcal{T}_D))? \tag{34}$$

If f is bounded then the answer is yes: simply express f in terms of indicator functions of local properties and apply Theorem 25. In general, (34) need not hold: for example, if $f(G, v)$ is the square of the degree of v then, since our assumptions give no control over $\sum_i d_i^2$, (34) can fail.

Suppose that $f(G, v)$ is the degree of v, so $\sum_{v \in C_1} f(C_1, v)$ is twice the number of edges in the giant component. Then, by (7), for any $\varepsilon > 0$ there is a C such that if $d_{\mathrm{conf}}(\mathbf{d}, D)$ is small enough, then

$$\sum_{v \in C_1 : d_{C_1}(v) \geqslant C} f(C_1, v) \leqslant \sum_{v \in G : d_G(v) \geqslant C} f(G, v) \leqslant \varepsilon n,$$

and considering the bounded function obtained by truncating f at C, we see that (34) holds in this case, even though f is unbounded. A similar argument can be applied to other unbounded f, leading to results concerning, for example, the number edges in the giant component between vertices of degree 2 and degree 3. We omit the details.

References

[1] E.A. Bender and R.E. Canfield, The asymptotic number of labeled graphs with given degree sequences, *J. Combinatorial Theory Ser. A* **24** (1978), 296–307.

[2] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *European J. Combin.* **1** (1980), 311–316.

[3] B. Bollobás, S. Janson and O. Riordan, The phase transition in inhomogeneous random graphs, *Random Struct. Algorithms* **31** (2007), 3–122.

[4] P. Erdős and A. Rényi, On the evolution of random graphs, *Magyar Tud. Akad. Mat. Kutató Int. Közl* **5** (1960), 17–61.

[5] N. Fountoulakis, Percolation on sparse random graphs with given degree sequence, *Internet Mathematics* **4** (2007), 329–356.

[6] A. Goerdt, The giant component threshold for random regular graphs with edge faults, *Theoret. Comput. Sci.* **259** (2001), 307–321.

[7] H. Hatami and M. Molloy, The scaling window for a random graph with a given degree sequence, *Random Struct. Algorithms* **41** (2012), 99–123.

[8] S. Janson, On percolation in random graphs with given vertex degrees, *Electron. J. Probab.* **14** (2009), 87–118.

[9] S. Janson and M.J. Luczak, A new approach to the giant component problem, *Random Struct. Algorithms* **34** (2009), 197–216.

[10] M. Kang and T.G. Seierstad, The critical phase for random graphs with a given degree sequence, *Combin. Probab. Comput.* **17** (2008), 67–86.

[11] M. Molloy and B. Reed, A critical point for random graphs with a given degree sequence, *Random Struct. Algorithms* **6** (1995), 161–179.

[12] M. Molloy and B. Reed, The size of the giant component of a random graph with a given degree sequence, *Combin. Probab. Comput.* **7** (1998), 295–305.

[13] B. Pittel, Edge percolation on a random regular graph of low degree, *Ann. Probab.* **36** (2008) 1359–1389.

[14] O. Riordan, The k-core and branching processes, *Combin. Probab. Comput.* **17** (2008), 111–136.

[15] O. Riordan, The phase transition in the configuration model, *Combin. Probab. Comput.* **21** (2012), 265–299.